



## Genome Dynamics of *Escherichia coli* during Antibiotic Treatment: Transfer, Loss, and Persistence of Genetic Elements In situ of the Infant Gut

Porse, Andreas; Gumpert, Heidi; Kubicek-Sutherland, Jessica Z.; Karami, Nahid; Adlerberth, Ingegerd; Wold, Agnes E.; Andersson, Dan I.; Sommer, Morten Otto Alexander

*Published in:*  
Frontiers in Cellular and Infection Microbiology

*Link to article, DOI:*  
[10.3389/fcimb.2017.00126](https://doi.org/10.3389/fcimb.2017.00126)

*Publication date:*  
2017

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Porse, A., Gumpert, H., Kubicek-Sutherland, J. Z., Karami, N., Adlerberth, I., Wold, A. E., Andersson, D. I., & Sommer, M. O. A. (2017). Genome Dynamics of *Escherichia coli* during Antibiotic Treatment: Transfer, Loss, and Persistence of Genetic Elements In situ of the Infant Gut. *Frontiers in Cellular and Infection Microbiology*, 7, [126]. <https://doi.org/10.3389/fcimb.2017.00126>

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



# Genome Dynamics of *Escherichia coli* during Antibiotic Treatment: Transfer, Loss, and Persistence of Genetic Elements *In situ* of the Infant Gut

Andreas Porse<sup>1†</sup>, Heidi Gumpert<sup>2†</sup>, Jessica Z. Kubicek-Sutherland<sup>3</sup>, Nahid Karami<sup>4</sup>, Ingegerd Adlerberth<sup>4</sup>, Agnes E. Wold<sup>4</sup>, Dan I. Andersson<sup>3</sup> and Morten O. A. Sommer<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Alfredo G. Torres,  
University of Texas Medical Branch,  
USA

### Reviewed by:

Swaine Chen,  
Genome Institute of Singapore,  
Singapore  
Timothy James Wells,  
The University of Queensland,  
Australia

### \*Correspondence:

Morten O. A. Sommer  
msom@bio.dtu.dk

<sup>†</sup>These authors have contributed  
equally to this work.

**Received:** 14 February 2017

**Accepted:** 28 March 2017

**Published:** 12 April 2017

### Citation:

Porse A, Gumpert H,  
Kubicek-Sutherland JZ, Karami N,  
Adlerberth I, Wold AE, Andersson DI  
and Sommer MOA (2017) Genome  
Dynamics of *Escherichia coli* during  
Antibiotic Treatment: Transfer, Loss,  
and Persistence of Genetic Elements  
*In situ* of the Infant Gut.  
Front. Cell. Infect. Microbiol. 7:126.  
doi: 10.3389/fcimb.2017.00126

<sup>1</sup> Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark, <sup>2</sup> Department of Clinical Microbiology, Hvidovre University Hospital, Hvidovre, Denmark, <sup>3</sup> Department of Medical Biochemistry and Microbiology, Uppsala University Biomedical Centre, Uppsala, Sweden, <sup>4</sup> Department of Infectious Diseases, University of Gothenburg, Sahlgrenska Academy, Gothenburg, Sweden

Elucidating the adaptive strategies and plasticity of bacterial genomes *in situ* is crucial for understanding the epidemiology and evolution of pathogens threatening human health. While much is known about the evolution of *Escherichia coli* in controlled laboratory environments, less effort has been made to elucidate the genome dynamics of *E. coli* in its native settings. Here, we follow the genome dynamics of co-existing *E. coli* lineages *in situ* of the infant gut during the first year of life. One *E. coli* lineage causes a urinary tract infection (UTI) and experiences several alterations of its genomic content during subsequent antibiotic treatment. Interestingly, all isolates of this uropathogenic *E. coli* strain carried a highly stable plasmid implicated in virulence of diverse pathogenic strains from all over the world. While virulence elements are certainly beneficial during infection scenarios, their role in gut colonization and pathogen persistence is poorly understood. We performed *in vivo* competitive fitness experiments to assess the role of this highly disseminated virulence plasmid in gut colonization, but found no evidence for a direct benefit of plasmid carriage. Through plasmid stability assays, we demonstrate that this plasmid is maintained in a parasitic manner, by strong first-line inheritance mechanisms, acting on the single-cell level, rather than providing a direct survival advantage in the gut. Investigating the ecology of endemic accessory genetic elements, in their pathogenic hosts and native environment, is of vital importance if we want to understand the evolution and persistence of highly virulent and drug resistant bacterial isolates.

**Keywords:** *Escherichia coli*, genome evolution, virulence plasmid dynamics, plasmid persistence, horizontal gene transfer, antibiotic treatment, urinary tract infections, infant gut

## INTRODUCTION

The human gut is home to a dense microbial ecosystem, the human gut microbiota, playing an important role in human health and physiology (Marchesi et al., 2015). As a commensal constituent of the gut microbiota in warm-blooded animals, *Escherichia coli* is highly adapted to the gut and colonizes the gastrointestinal tract within the first hours of life (Drasar and Hill, 1974). However, some environmental and commensal *E. coli* isolates have acquired genetic factors that allow them to cause disease within the digestive tract or when transferred to other body sites such as the blood, brain, and urinary tract (Smith et al., 2007). While diarrheagenic *E. coli* are a common cause of gastro intestinal infections in third world countries and travelers, extraintestinal pathogenic *E. coli* (ExPEC) are facultative pathogens that reside in the human gut microbiota but occasionally establish in extra-intestinal body sites (Köhler and Dobrindt, 2011). Here, urinary tract infections caused by ExPEC are among the most common bacterial infections in developed countries, where patients are often infected via transmission of strains from their own intestinal flora to their urinary tract (Foxman, 2010).

The broad adaptation of *E. coli* to the gut environment and extraintestinal body sites is reflected in the remarkable genetic diversity within the species. This genetic flexibility is largely facilitated by horizontal gene transfer (HGT) of accessory genetic elements including plasmids and phages (Brzuszkiewicz et al., 2009). These elements are widely present within the gut microbiota and can provide their bacterial hosts with antibiotic resistance or virulence factors (Salys et al., 2004; Sommer et al., 2010). Acquiring virulence genes might not only influence the risk and severity of infections caused by the pathogen, but has also been suggested to assist in the general persistence of commensal bacterial strains of the gut (Diard et al., 2010; Chen et al., 2013). Indeed, virulence determinants such as those involved in adhesion, biofilm formation and iron acquisition correlate with prolonged colonization in the digestive tract (Adlerberth et al., 1998; Nowrouzian et al., 2003).

Recent studies into the dynamics of clinical bacterial genomes at genomic resolution have been carried out with time-series sampling and underlines the high plasticity of plasmids and their host associations *in situ* (Conlan et al., 2014, 2016). Conjugative plasmids are of particular interest, as they are the main vehicles of HGT in *E. coli*, playing an essential role in the adaptation toward antibiotics or specific host niches (Johnson and Nolan, 2009; Norman et al., 2009).

Whereas, much effort has been devoted to study the survival conditions of plasmids *in vitro* (Slater et al., 2008) our knowledge on the behavior of plasmids *in situ* of their native hosts and natural environment is limited (Karami et al., 2007; Conlan et al., 2014, 2016). In order for a plasmid to persist in the long term, it needs to either be stably segregated upon cell division, confer a fitness advantage to its host, or transfer at high enough rates to compensate the lack of the latter two (Simonsen, 1991; Slater et al., 2008). As most plasmids do not exhibit sufficient rates of transfer to survive without selection, stable inheritance, and

adaptive traits are key to their long term survival (Simonsen, 1991).

To elucidate the genome dynamics of *E. coli* in its native environment of the gut, we genome sequenced individual *E. coli* isolates over the first year of an infant's life. We conduct *in vitro* and *in vivo* competition assays to elucidate the selective drivers of the observed dynamics, and gain a deeper understanding of the endemic mobile elements contributing to the dissemination of virulence and antibiotic resistance factors.

## MATERIALS AND METHODS

### Genome Sequencing of *E. coli* Lineages

The strains were isolated and typed as part of a previous study by Karami et al. (2007). These were cultured in LB broth and genomic DNA was isolated using an UltraClean<sup>®</sup> Microbial DNA Isolation Kit (MoBio Laboratories, Inc., California). Sequencing libraries were prepared using the TruSeq and Nextera XT (Illumina, California) protocols. Illumina HiSeq sequencing was performed by Partners HealthCare Center for Personalized Genetic Medicine (Cambridge, Massachusetts).

### Sequence Analysis

Genomes for each sequenced isolate were assembled using Velvet (v1.2.10; Zerbino and Birney, 2008) and annotated via RAST (Aziz et al., 2008). Reads from the isolates were mapped onto the reference, e.g., earliest isolated, genome via Bowtie2 (2.1.0; Langmead et al., 2009), and single nucleotide polymorphisms (SNPs) were enumerated via SAMTools (0.1.19; Li et al., 2009). The SNP threshold was set to include SNPs with a phred score of above 30 and at least 90% of the high-quality reads at the site as the variant. Additionally, to ensure that all isolates within a lineage consisted of the same genomic content as the representative isolates, genomic areas lacking mapped read coverage were identified using BEDTools (2.18.2; Quinlan and Hall, 2010).

The pNK29 plasmid was assembled into a circular plasmid with aid from plasmid alignments produced using MUMer (Kurtz et al., 2004). Contigs belonging to the pNK29 antibiotic resistance plasmid were first identified in lineage B as the new genetic material of the isolate at 32 days, and then used to identify the corresponding contigs in the lineage A genome. The RAST annotations for this plasmid were refined based on homologous genes in pOLA52 (NC\_010378.1) that were either missing or incorrect in pNK29.

### Plasmid Identification and Comparison

Other plasmids were identified by first separating contigs based coverage to infer copy-number relative to genomic contigs and then by grouping contigs together with similar abundances. The average coverage of each contig was determined using BEDTools (Quinlan and Hall, 2010). Plasmid incompatibility grouping was done using the PlasmidFinder tool (Carattoli et al., 2014). Homologous previously sequenced plasmids were identified using BLAST and the NCBI nt database (Altschul et al., 1990). Circular plasmid diagrams were created using the BLAST ring image generator (BRIG; Alikhan et al., 2011).

For pNK29-2, blastn searches of the plasmid contigs revealed 14 plasmids with very high identity (99%) and hits with a pNK29-2 coverage of >97% were selected (**Figure 3** and **Table S4**). As an exception, pECO-bc6 was also included despite its lower coverage (88%) to illustrate deletion of plasmid accessory genes flanked by inverted repeats. The EasyFigure software was used for linear comparison of plasmid sequences displayed in **Figure 4** (Sullivan et al., 2011).

The core genome of the *E. coli* hosts listed in **Figure 3** was estimated using ROARY via annotations from PROKKA (Seemann, 2014; Page et al., 2015). The aligned, ungapped core genome was used to construct a maximum likelihood phylogenetic tree using the RAxML software (Stamatakis, 2014). MLST types were assigned using MLSTfinder (Larsen et al., 2012), and *fimH* types were assigned using the sequences referred to by Dias et al. (2010).

## Strain Tagging and pNK29-2 Plasmid Curing

Lineage A and B strains isolated at the first time point were tagged with antibiotic resistance markers to allow quantification during competitive fitness experiments, plasmid loss experiments, assessment of conjugation ability, and plasmid curing. Resistance cassettes conferring resistance to Chloramphenicol and Kanamycin respectively were amplified from cloning vectors of the pZ system (Lutz and Bujard, 1997) and inserted into the chromosomal *araB* gene of the Lineage A and B strains using the Lambda Red recombineering system of pTKRED (Kuhlman and Cox, 2010). The following regions of homology were used for insertions into *araB*: 5'-GTAGCGAGGTTAAGATCGGTAATC ACCCCTTTCAGGCGTTGGTTAGCGTT-3' and 5'-GCCTAACGCACTGGTAAAAAGTTATCGGTACTTCCACCTGCGACA TTCTGA-3'.

The pNK29-2 plasmid was tagged with the Sh *ble* Zeocin resistance gene in a transposase gene located at 61 kb using the following homology ends: 5'-CTTCGGGAACGCTGTAACGAT TACCACCAACCTCGATATAGCTGTCCCGG-3' and 5'-TAA CAACGGGAAGTCGTGTTCAACTCCGATTCCCTGTTGC TGCCCGACC-3'.

To cure pNK29-2 we disrupted the *stbA* gene of the *stbAB* stability operon with a Kanamycin resistance marker using the following homology ends: 5'-CATAAATGTGATGTGT GAAGTATGATGATATTTTGACACGGTAACCTGAGTA GGGATAACAGGGTAAT-3' and 5'-TTCATTTTAAGAC GCACATCATTGCTCCTGCAACGAATCAGTAG CTAGGGATAACAGGGTAAT-3'. These contain the 18 bp I-SceI endonuclease site enabling *in situ* double digestion of the recombinant plasmid to enhance plasmid loss via induction of the I-SceI endonuclease from pTKRED. Following recombineering, recombinants were selected on Kanamycin containing plates incubated at 30°C to retain the pTKRED plasmid. These were verified with PCR to confirm insertion into the *stbA* gene. Verified colonies were inoculated directly into LB medium containing 0.5% w/v L-arabinose and grown at 30°C for 2 h to induce expression of the pTKRED encoded I-SceI endonuclease and subsequently switched to 42°C for

3 h to allow curing of the temperature sensitive pTKRED vector. The culture was diluted and plated on LB to obtain single colonies that were confirmed for plasmid curing by the absence of growth on Kanamycin (pNK29-2) and Spectinomycin (pTKRED). The curing of pNK29-2 was validated by PCR using primers targeting two distinct loci of the pNK29-2 backbone: Upstream *stb*: 5'-CTCAACAAGGGTTATTGC-3', downstream *stb*: 5'-GAATGGCAAATGAAACG-3' and Upstream 61 kb transposase: 5'-GAATGGCAAATGAAACG-3', downstream 61 kb transposase: 5'-AGAAGGCTGCGGTGCTGAAG-3'. The plasmid-cured variant of the lineage A strain was re-transformed with the Zeocin tagged pNK29-2 plasmid to control for potential effects of the curing process.

## Conjugative Transfer Assessment

In order to test the ability of pNK29-2 to conjugate, outgrown over-night (O/N) cultures as well as exponentially growing cultures of the lineage A strain and *E. coli* MG1655::tetA was mixed equally and incubated O/N. Incubations were done at 37°C and 30°C on a solid agar surface as well as in liquid cultures without shaking. Additional tagging at the 32 and 26 kb positions were carried out to ensure that the initial insertion at the 61 kb position of pNK29-2 was not the cause of the dysfunctional conjugation ability.

## In vitro Competition, Growth Rate, and Plasmid Loss Assays

Two O/N cultures were diluted to the same OD and mixed 1:1 in LB medium. The competition was carried out in 1.5 ml cultures and a volume of 1.5 µl was transferred to a fresh well every 20 h. From OD measurements the number of generations was estimated to ~10 generations/transfer. The ratio of the competitors was determined as the fraction of colonies on Chloramphenicol agar plates compared to plates containing Kanamycin. Plasmid loss was assessed by comparing plate-counts of at least 100 colonies on LB and Zeocin agar plates. In addition, colonies from LB plates were streaked on Zeocin containing plates and PCR (using the primers listed in the plasmid curing section) was performed on a subset of colonies to verify plasmid presence. Competitions in iron-limited medium was carried out as for the LB competitions described above, except that M9 medium [M9 minimal medium (standard), 2010] with 5 µM FeCl<sub>3</sub> was used instead of LB. OD measurements of growth rates in M9 5 µM FeCl<sub>3</sub> were conducted in 96-well plates containing 150 µl medium/well using a ELx808 plate reader (BioTek, USA). Breathe-Easy (Sigma-Aldrich) film was applied to minimize evaporation during measurements. OD at 600 nms was measured with 5-min intervals for 24 h and incubated with shaking at 37°C between measurements.

## In vivo Competition Experiments

Female BALB/c mice (5–6 weeks old) were used in all *in vivo* studies (Charles River Laboratories, distributed by Scanbur). All mice were pre-treated orally with streptomycin as described previously (Lasaro et al., 2014). Briefly, streptomycin sulfate salt (Sigma-Aldrich) was added to the drinking water at 5 g/L, along with 5 g/L of glucose, to enhance taste, for 72 h followed by 24



h of fresh water (no drug or glucose) to allow the streptomycin to clear the animal's system prior to infection. A single colony of each *E. coli* strain was grown in LB shaking overnight at 37°C. Cells were pelleted, washed once in PBS and re-suspended in PBS (13 mM phosphate with 137 mM NaCl at pH 7.4), and then mixed in a 1:1 ratio of *E. coli* lineage A isolate with pNK29-2 to the cured variant. Fifteen mice that were pre-treated with streptomycin followed by 24 h without drug were administered 100 µL containing  $2 \times 10^8$  CFU of this 1:1 *E. coli* mixture by oral gavage. Feces was collected at days 2, 4, 7, and 12 post-infection. Additionally, on day 12 following termination of the experiment, a segment of the small intestine was removed. The feces and small intestine segment were homogenized in PBS, serially diluted, and equal amounts were plated on LA-Cam (25 µg/ml chloramphenicol, selecting for the pNK29-2 containing strain) and LA-Kan (50 µg/ml kanamycin, selecting for the cured strain). Following overnight incubation at 37°C, CFUs were enumerated and subsequently replica plated from both LA-Cam and LA-Kan to LA-Zeo (40 µg/ml zeocin) to screen for the presence and transfer of the pNK29-2 plasmid containing a Zeocin-resistance marker. CFU values were normalized per gram of tissue (CFU/g). The competitive index was calculated by dividing the OUTPUT on days 2, 4, 7, and 12 (CamR CFU/g divided by KanR CFU/g) by the INPUT on day 0 (CamR CFU/g divided by KanR CFU/g). The input value was 1.04 indicating a 1:1.04 initial ratio of *E. coli* lineage A isolate with pNK29-2 relative to the cured *E. coli* lineage A isolate. The non-parametric Mann-Whitney *U*-Test was used to compare the sample populations. The *P*-values indicate the probability of falsely rejecting the null-hypothesis of equal population means.

## Ethics Statement

Animal experiments were performed in accordance with national (regulation SJVFS 2012:26) and institutional guidelines. The Uppsala Animal Experiments Ethics Review Board in Uppsala, Sweden approved all mouse protocols undertaken in this study under reference no. 154/14. Animal experiments were performed at the Swedish National Veterinary Institute (SVA) in Uppsala, Sweden.

## RESULTS

The current study material was obtained from an infant enrolled in the ALLERGYFLORA study with the original purpose of correlating the composition of the gut microbiota to the development of allergies later in life (Adlerberth et al., 2006).

This infant was subjected to long-term antibiotic treatment as a consequence of a urinary tract infection, and was selected for this study due to an observed change in the resistance profile of *E. coli* strains isolated from this infant (Karami et al., 2007).

Fecal samples were obtained from the infant at 2, 9, 16, and 32 days, and 2, 6, and 12 months after birth (Figure 1). *E. coli* lineages were identified by morphological and biochemical characteristics as well as subsequent confirmation by PFGE and random amplified polymorphic DNA (RAPD) typing. The two main lineages were designated "A" and "B." *E. coli* lineage A was recovered at all the sampling time points and lineage B was only transiently present in the samples collected from day 9 to 32

days of age (Karami et al., 2007; Figure 1). At 11 days of age, a UTI infection was diagnosed and trimethoprim administered intravenously (i.v.) for 5 days. However, due to the subsequent presence of enterococci in addition to *E. coli*, the antibiotic treatment was changed to i.v. ampicillin for 5 days followed by oral amoxicillin treatment for an additional 8 days. Lastly, the infant was administered trimethoprim prophylactically for the following 7 months (Figure 1).

## Transfer of an Antibiotic Resistance Plasmid between Two Distinct *E. coli* Lineages Co-Colonizing the Infant Gut

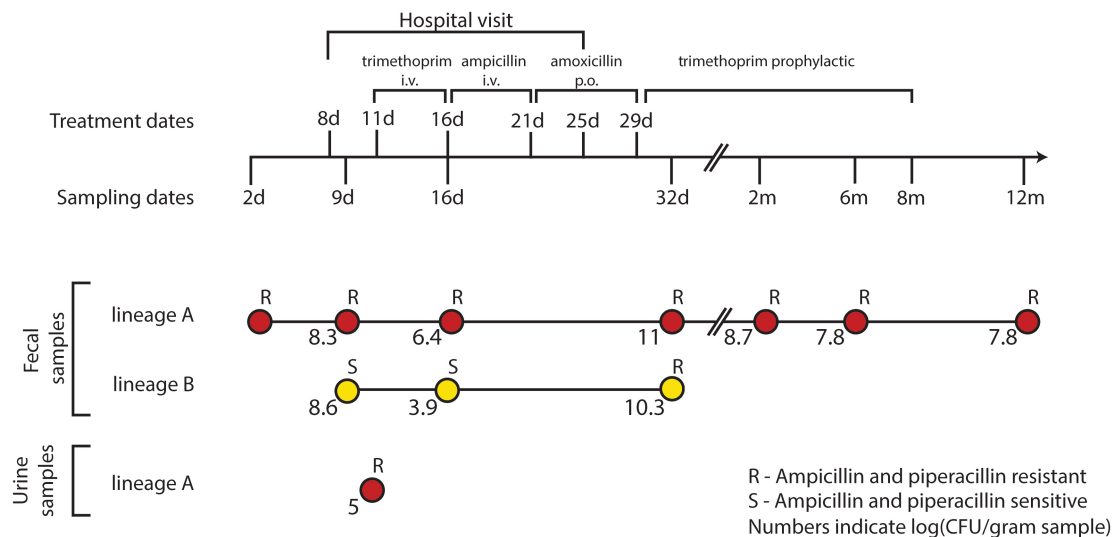
We sequenced the genomes of the *E. coli* isolates, obtained from the fecal and urine samples of the infant, which confirmed the lineages previously identified via RAPD and PFGE typing. Comparing the genomic similarity of the two lineages revealed that the lineages did indeed originate from two different strains; with the initial isolate of lineage A (4.91 Mb) sharing only 77% of lineage B's (5.45 Mb) genomic content.

To assess the genomic divergence of the individual lineage A isolates, during the sampling period, a SNP-based phylogenetic tree was constructed (Figure 2, Table S1). Only one SNP was found when comparing the genomes of lineage A isolates collected at 2, 9, and 16 days to the UTI isolate (Figure 2). Given the very high sequence similarities between these isolates, lineage A colonizing the gut microbiota was assumed to be the cause of the UTI.

From the annotated genomes of lineage A and lineage B isolates, we identified several factors that could contribute to the pathogenicity of these strains. The genome of the uropathogenic lineage A encoded the type 1 fimbriae FimH among other adhesion factors (*AidA-I* and *yqi* encoded adhesions), siderophore (enterobactin and yersiniabactin) transporters and hemin receptors (TonB-system) as well as enterotoxins (*senB* and *vat*) and Hemolysin E. Although the lineage B isolates did not cause infection, its genomic content reveals similar virulence factors such as the type 1 fimbriae (*fimH*), serum survival factors (*iss*) and iron acquisition (aerobactin synthesis and transport), but no enterotoxins.

While no SNPs were detected in the three isolates from lineage B, we report the sequence of pNK29, a 42.2 kb TEM-1b encoding conjugative IncX plasmid, that transferred from lineage A to lineage B *in situ* of the gut (Karami et al., 2007; Figure S1, Table S2). This plasmid was first detected in lineage B at 32 days, and coincided with high resistance toward ampicillin compared to earlier isolates (Figure 1). Similar conjugative plasmids of the IncX family are prevalent in pathogenic *E. coli*, as well as other Enterobacteriaceae isolated from humans and animals, playing an important role in the dissemination of antibiotic resistance genes (Norman et al., 2008; Toro et al., 2014).

Karami et al. reported an increase in lineage A counts from  $10^{6.4}$  CFU/g fecal matter to a density of  $10^{11}$  CFU/g as the infant was switched from trimethoprim to ampicillin and amoxicillin treatment during the UTI infection from day 16 to 32 (Karami et al., 2007). Such events can increase population size, and thus the probability of plasmid transfer and enrichment of pNK29



**FIGURE 1 | Resistance profiles of co-existing *E. coli* lineages during the course of a urinary tract infection and antibiotic treatment.** *E. coli* isolates were obtained during the first 12 months of an infant's life. At 8 days of age, the infant was admitted to the hospital due to a urinary tract infection from which the lineage A clone was isolated at day 11. After 5 days of i.v. trimethoprim the treatment was switched to 5 days of i.v. ampicillin followed by an additional 8 days of oral (p.o.) amoxicillin. Trimethoprim was administered to prevent reoccurring infections for the following 7 months. Along the course of treatment, lineage B acquired a TEM-1b encoding IncX plasmid from lineage A; rendering both lineages resistant to the  $\beta$ -lactam treatment at day 32.

bearing cells. A similar increase in population counts from  $10^{3.9}$  to  $10^{10.3}$  CFU/g was observed for lineage B as a result of pNK29 acquisition and antibiotic selection (Figure 1).

Interestingly, pNK29-bearing lineage B isolates were no longer detected in the subsequent samples collected after cessation of amoxicillin treatment (Figure 1). Plasmids often impose a fitness cost upon first encounter with new host backgrounds which could render lineage B less fit in the absence of selection (Porse et al., 2016). Measuring the *in vitro* competitive fitness of the initial plasmid-carrying lineage B isolate revealed a burden of carriage ( $-4.9\%$ ,  $sd \pm 4.1\%$ ,  $P = 0.046$ ); indicating that a counterselection of lineage B, due to plasmid invasion, might have taken place after discontinuation of amoxicillin treatment.

## Major Genomic Events of Lineage a during Gut Colonization

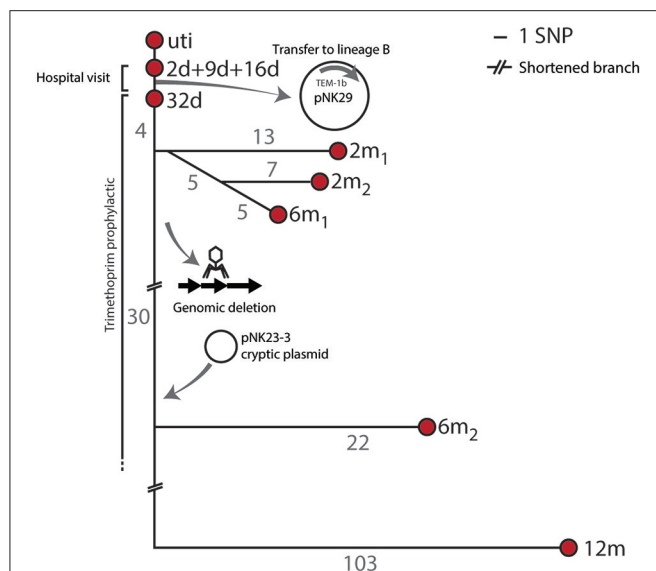
Apart from the pNK29 plasmid, all isolates of lineage A carried a large virulence plasmid, designated pNK29-2, which was detected throughout the year of sampling. In addition to these two large plasmids, a novel plasmid-element was detected in the 6 m<sub>2</sub> and 12 m isolates (Figure 2). This small (2,545 bp) cryptic plasmid was termed pNK29-3 and had a low GC content of 33.4%. Two open reading frames were identified on the plasmid, which encode putative mobilization and replication proteins (Figure S2). By comparing the coverage depth of the plasmid to the average coverage depth of the genome, we estimate the copy number to be around nine plasmids per cell. BLAST-analysis revealed a high resemblance to the pIGMS31 previously isolated from *Klebsiella pneumoniae* as well as pEA1 (DQ659147.1) isolated from a Brazilian *Pantoea agglomerans* strain (Figure S2; Smorawinska et al., 2012; Carattoli et al.,

2014). It was shown that while pIGMS31 can be mobilized to Alpha- and Gamma-proteobacteria, it replicates via a rolling circle mechanism that functions in Gammaproteobacteria only (Smorawinska et al., 2012). Therefore, pNK29-3 likely originates from other Gammaproteobacteria constituents of the gut flora.

Coinciding with the acquisition of pNK29-3 by lineage A, the 6 m<sub>2</sub> and 12 m isolates of lineage A were also missing ~54 kb of their chromosome compared to the previous isolates. Annotations and flanking attR and attL sites of this region suggested that it was an integrated phage. Using the PHAST server and BLAST searches against NCBI GenBank, we detected a high similarity to the Siphoviridae prophage of *E. coli* strains FH199 and UMNK88 (Zhou et al., 2011). Losing this prophage did not alter the resistance profile of the strain, but might have been a result of negative selection imposed by increased excision activity as a consequence of the cellular stress imposed by antibiotic exposure (Beaber et al., 2004).

## Lineage a Harbored a Highly Disseminated and Stable Virulence Plasmid

The pNK29-2 plasmids of lineage A displayed very high sequence identity to several widely disseminated plasmids deposited in NCBI's Genbank (Figure 3). Interestingly, 12 highly similar plasmids were previously isolated from various pathogenic *E. coli* strains and one originated from *Klebsiella pneumoniae*. In particular, the endemic pUTI89 plasmid was found to have only 7 SNP differences to the pNK29-2 plasmid, and aligning the contigs from this plasmid showed that there were no additional insertions (Chen et al., 2006). Only two of the SNPs lead to non-synonymous changes. These are located in the *rsvB* gene, a resolvase, and in the *traE* gene, a conjugal transfer protein for F



**FIGURE 2 | SNP tree and genomic events of lineage A.** Branch-numbers show the amount of SNPs separating the isolates obtained from 2 days to 12 months after birth. The majority of samples were isolated during hospitalization and prophylactic treatment. No SNPs were detected between the isolates collected at days 2, 9, and 16, and only one SNP (in the blaTEM1b promoter) was detected between these isolates and the isolate collected at 32 days. Two isolates were included at the 2 months sampling point (2 m1 and 2 m2) and these shared 5 SNPs in common compared to the 32 days isolate. Two 6 month isolates were also included. The infant was subjected to antibiotic treatment from day 11 to 8 months after birth. Apart from SNP differences one sub-lineage had undergone major genomic changes by means of chromosomal deletions and acquisition of new plasmid DNA.

pilus assembly (Table S3). We tested the ability of pNK29-2 to conjugate to *E. coli* MG1655, in liquid and on solid media, and we did not detect any transconjugants; implying that the mutated *traE* transfer gene is dysfunctional in pNK29-2.

## pNK29-2 Resembles Virulence Plasmids Found in a Diverse Set of Pathogenic *E. coli*

The *E. coli* strain UTI89, harboring the pUTI89 plasmid, is an archetypical uropathogenic *E. coli* (UPEC) strain isolated from a patient with an acute bladder infection (Mulvey et al., 2001). The pUTI89 plasmid belongs to the IncFIB/IIA incompatibility group and shares several characteristics with the F-plasmid including a full *tra* operon for conjugative transfer (Chen et al., 2006). Additionally, its core backbone includes stability mediating genes such as the *ccdA-ccdB* toxin-antitoxin system and the *stbAB* operon ensuring stable inheritance upon cell division (Cusumano et al., 2010). Several of the lesser conserved plasmid regions can be related to virulence and overall adaptation to the human host. These encode the enterotoxigenicity (*senB*), copper tolerance (*scsC/scsD*) and iron acquisition factors. The *cjrABC* operon encodes proteins involved in iron transport that also cause sensitivity to colicin and has shown involvement in UTI virulence (Smajcs and Weinstock, 2001; Cusumano et al., 2010).

While a majority of the bacterial hosts carrying these virulence plasmids is associated with UTIs, extremely similar plasmids have been isolated across *E. coli* patho- and sequencetypes from all over the world (Figure 3). For example, the prototype neonatal meningitis *E. coli* strain RS218, isolated in 1974, carries a virulence plasmid virtually identical to pUTI89 and pNK29-2, which has been shown to play an important role in its pathogenicity (Wijetunge et al., 2014).

Likewise, the genomic backgrounds hosting these plasmids are diverse (Figure 3) and include strains of the dominating extraintestinal pathogenic *E. coli* ST131 of major clinical importance (Stoesser et al., 2016). The geographical locations where these strains have been isolated varies, with highly similar plasmids isolated from different strains in the US, Japan, Canada, and Sweden, suggesting that these pUTI89-like plasmids are globally disseminated in a non-clonal fashion (Figure 3).

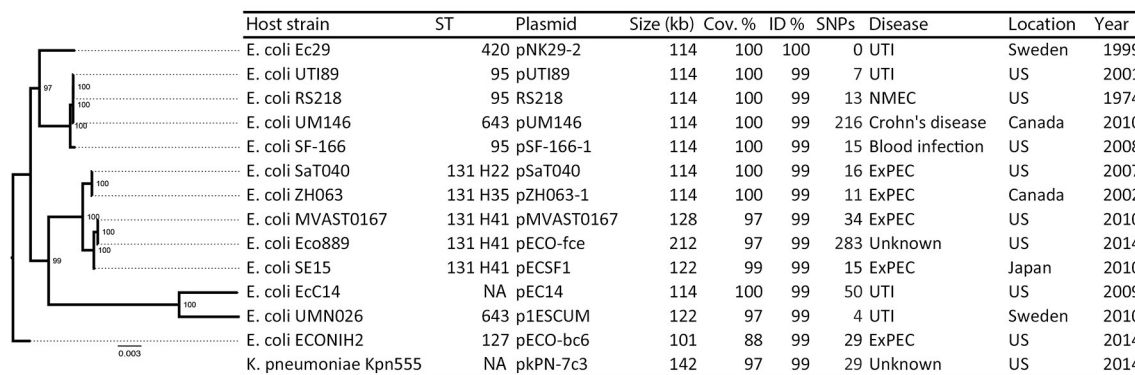
## Plasmids Similar to pNK29-2 Carry Antibiotic Resistance Genes Obtained from Independent Insertion Events

Although these plasmids share substantial homology, some show variation within defined but variable “genetic load” regions of their plasmid backbone (Figure 4).

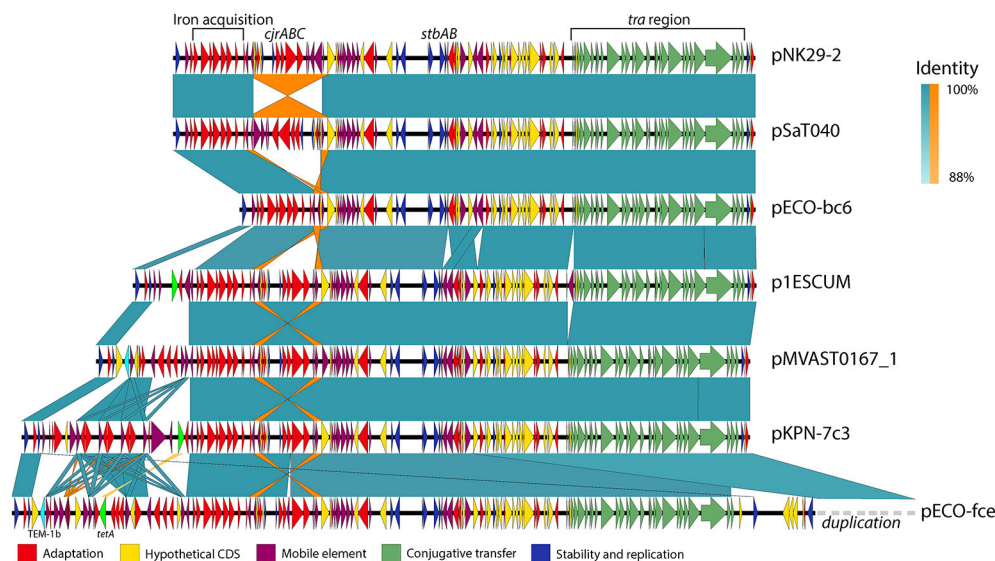
Antibiotic use is a strong selection force and some of the plasmids similar to pNK29-2 encode one or more, antibiotic resistance genes organized among mobile element associated cassettes. For instance, the pESCUM and pKPN-7c3 plasmids have acquired the *tetA* gene, conferring tetracycline resistance, at different positions within their genetic cargo region, but associated with the same mobile elements (Figure 4; Johnson et al., 2016). pMVAST0167\_1 and pECO-fce encode multiple antibiotic resistance genes from their genetic load region and share a TEM-1b gene in the same location and accommodate a similar integron with genes conferring resistance toward aminoglycosides (*aadA5*), sulphonamides (*sul1*) and trimethoprim (*dfrA17*). Compared to pMVAST0167\_1, pECO-fce encodes several additional resistance genes (*sul2*, *tetA*, *strA*, *strB*, and *tnrB*) from a cassette inserted between the TEM-1b gene and the *int1* integron (Figure 4).

## pNK29-2 Confers a Fitness Cost to its Host *In vitro*

While the pNK29 plasmid, encoding the TEM-1b  $\beta$ -lactamase, was strongly selected for during the  $\beta$ -lactam treatment administered here (Karami et al., 2007), it is not obvious how large virulence plasmids persist in the gut. The lineage A strain persistently colonized the gut for the entire duration of the study, from 2 days to 1 year after birth, suggesting that it is well adapted to the human host. Such long term survival could be supported by general persistence factors of UPEC strains, suggesting an overlap between UTI virulence and gut persistence factors (Nowrouzian et al., 2005; Chen et al., 2013). Only minor selection is required for a plasmid to survive if it does not impose a significant fitness cost on its host. We used antibiotic resistance markers to tag the plasmid and the host chromosome of lineage A in order to separate the plasmid-bearing variant from the cured variant and



**FIGURE 3 | Overview of plasmids resembling pNK29-2 and their host phylogeny.** Thirteen plasmids with high similarity to pNK29-2 were downloaded from Genbank (see Table S4 for accession numbers and references). Info on host strain, its disease associations and geographical origin was obtained from the literature. The “Year” column denotes the first mentioning of the host strain in the literature unless the isolation year was clearly stated. *In silico* MLST typing was performed and FimH types were added to the ST131 clade highlighting their internal diversity. A core-genome based maximum-likelihood tree was constructed to illustrate the diversity of the *E. coli* plasmid hosts. Node numbers are bootstrap confidence values.



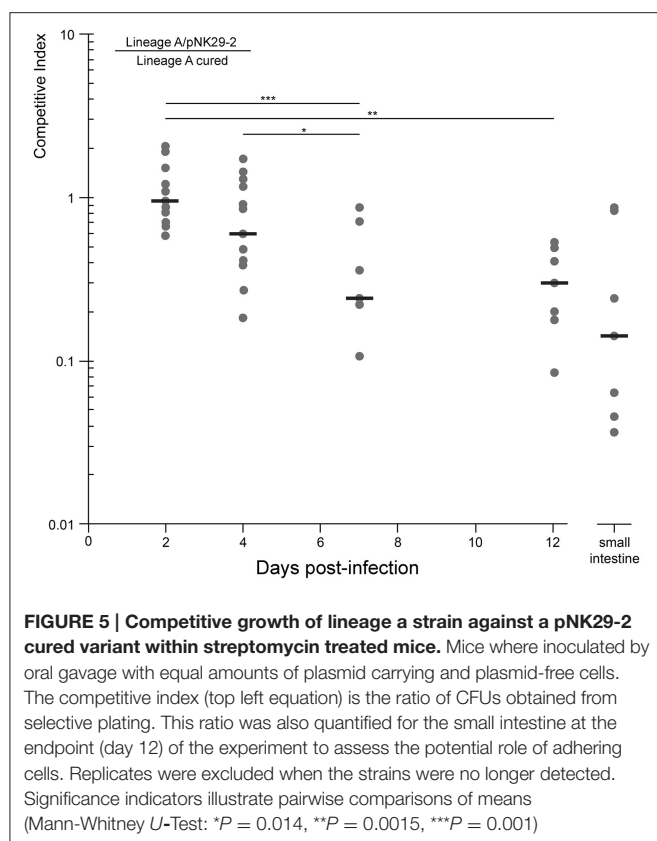
**FIGURE 4 | Genetic variability and conservation within similar virulence plasmid backbones.** six out of the Thirteen virulence plasmids similar to pNK29-2 showed signs of major restructuring events and are illustrated here. While these plasmids show some degree of variation, they also share a conserved core of transfer (green), stability (blue), and virulence genes (red). Mobile elements constituting inverted repeats (highlighted in orange) allow for instability of the *cjrABC*-containing region exemplified in pSaT040 and pECO-bc6. Additional genes involved in antibiotic resistance have been inserted upstream the fully conserved iron acquisition cluster of the genetic load region in p1ESCUM, pMVA0167\_1, pKPN7c3, and pECO-fce. The TEM-1b and *tetA* genes are highlighted in cyan and light green respectively. pECO-fce is significantly larger than the remaining plasmids due to a duplication of its transfer region which has been condensed for simplicity. Colored shades illustrate BLAST identity of pairwise comparisons and orange shades highlight inverted regions.

detect plasmid loss as well as transfer events. We used markers previously used for tagging in similar fitness experiments, where they did not impose a measurable cost (Chen et al., 2013). Similarly, we also confirmed that the *Sh ble* Zeocin resistance gene, used for plasmid tagging, did not impose a significant cost during 5 days of competitive growth against the non-tagged variant (two sample *t*-test,  $P = 0.11$ ).

The cured and plasmid bearing strains were mixed in equal proportions and propagated in LB medium for 8 days without

selection. From CFU quantifications on selective agar plates, the average *in vitro* fitness cost of the plasmid-carrying strain was measured to be  $0.92 \text{ sd} \pm 0.25\%$  per generation. Although this cost of plasmid carriage is low, one would expect a steady decline of the plasmid bearing cells under these growth conditions. Due to the presence of putative iron acquisition systems on the pNK29-2 plasmid, we also tested the competitive fitness as well as absolute growth rates in iron-limited M9 minimal medium containing  $5 \mu\text{M}$   $\text{FeCl}_3$  mimicking the lower range of





physiological concentrations (Wang, 1996). We were not able to detect a significant advantage in terms of growth rate (Two sample *t*-test,  $P = 0.43$ —Figure S3) or competitive fitness of the plasmid carrying strain when mixed 1:1 in the same medium ( $1.07 \pm 2\%$ , one sample *t*-test,  $P = 0.45$ ).

## The Plasmid Carrying Strain is Outcompeted *In vivo* of the Mouse Gut

Because the plasmid encodes a diverse set of factors thought to be involved in iron acquisition, toxin production as well as several hypothetical proteins, that might be selected for in more complex *in vivo* settings, we set out to test the competitive fitness in a gut environment of streptomycin treated mice. The plasmid-cured strain was competed against its plasmid-carrying ancestor in the mouse gut and we measured the proportion of plasmid-carrying to plasmid-free cells in the feces over the course of 12 days (Figure 5). Here, we observed a steady decline in plasmid-carrying cells compared to plasmid-free cells with the plasmid-free cells dominating the average population after 7 days of direct competition. Interestingly, the plasmid-carrying populations avoid complete out-competition by plasmid-free cells in this time-span (Figure 5).

To investigate whether the measurements obtained from feces were representative of the intestinal contents, including cells adhering to the intestinal wall, we measured the proportion of plasmid-bearing cells sampled directly from the small intestine of the mice at the end of the 12 days experiment. There was no significant difference between the direct sampling compared to

fecal counts (Mann-Whitney *U*-Test,  $P = 0.27$ ) and >10% of the cells contained the plasmid in the average replicate population at this stage (Figure 5).

From the competition experiment we calculated the average fitness cost per generation of the plasmid in the mouse gut to be  $0.83 \text{ sd} \pm 0.2\%$ , assuming that the competing *E. coli* underwent 10 generations per day (Rang et al., 1999; Lee et al., 2010; Myhrvold et al., 2015). This cost is slightly lower, but not significantly different, from the plasmid cost measured *in vitro* (Two sample *t*-test:  $P = 0.61$ ).

## pNK29-2 is Stably Inherited Despite its Cost

Although the plasmid-bearing strain is less fit than the plasmid-cured variant *in vivo*, it seems possible that a minor subpopulation of plasmid-bearing cells can persist for extended time periods; especially if competition from plasmid-free daughter cells is postponed via stable inheritance mechanisms. As for all the virulence plasmids analyzed here (Figure 3), pNK29-2 carries the highly conserved *stbAB* stability operon (Figure 4) that encodes active segregation machinery; ensuring that the plasmid is stably segregated to both daughter cells during cell division (Guynet et al., 2011). To test the segregational stability of pNK29-2 in its native lineage A host (day 0 isolate), we conducted 14 days of serial passaging in LB medium, corresponding to ~140 generations of growth. In such a setup, plasmid-free segregants would eventually take over the population due to the cost of plasmid carriage. However, we did not observe any plasmid-free cells within this time span with a detection limit of plasmid-free cells of ~1%. Similarly, no plasmid-free cells were detected in the *in vivo* competition assays, during the 12 days of gut colonization.

Taken together, these results imply that virulence plasmids, such as pNK29-2, have no direct advantage in gut colonization but are able to persist in spite of a small, but significant, fitness cost due to efficient plasmid inheritance mechanisms.

## DISCUSSION

Culture independent methods based on metagenomic sequencing have been used to investigate the abundance profiles of strains colonizing the gut over time (Morowitz et al., 2011; Brown et al., 2013; Sharon et al., 2013). However, these methods are limited in their ability to observe genomic events at high resolution such as horizontal transfer and single nucleotide variations.

Due to our longitudinal sampling and the high resolution of single isolate genome sequencing, we were able to observe a glimpse of the complex genome dynamics of *E. coli* in its native settings. We confirm a gut-inhabiting strain as the origin of a bladder infection; supporting the general belief that UTIs are caused by gut inhabiting *E. coli* strains that eventually enters the urethra (Chen et al., 2013). Furthermore, we report the sequence of pNK29, a novel 42.2 kb IncX plasmid carrying a TEM-1b  $\beta$ -lactamase, which was transferred between the two co-existing *E. coli* lineages of the gut. Few phenotypic reports exist documenting plasmid mediated HGT of antibiotic resistance genes between bacteria in the human gut and our data supports

that transfer of resistance genes take place in the gut, and may be enhanced by antibiotic treatment (Bidet et al., 2005; Karami et al., 2007; Trobos et al., 2009; Goren et al., 2010). The recipient lineage was only sampled at one time-point after the termination of  $\beta$ -lactam antibiotic administration (day 32) and declined to undetectable levels thereafter. This could indicate a negative selection of pNK29-carrying lineage B isolates in the absence of antibiotic selection; however, confirming the role of pNK29 in the counterselection of lineage B in the gut will require further *in vivo* competition experiments. Lineage B showed a large drop in population counts when subjected to the first round of trimethoprim treatment upon hospitalization (**Figure 1**). The prophylactic administration of trimethoprim coincided with the disappearance of lineage B, and could be another likely explanation for its absence in the consecutive time points.

While we did not observe any genomic alterations of lineage B apart from the acquisition of pNK29, lineage A experienced chromosomal deletions and acquired a cryptic plasmid as well as a high number of SNPs during the sampling period (**Figure 2**).

Genome plasticity is believed to play a crucial role in the adaptation of pathogens to the selective forces imposed by the immune system or the remaining microbiota within a human host (Brzuszkiewicz et al., 2009). The mutation rate observed for lineage A was high and resemble that of mutator phenotypes that are often enriched among UPEC isolates (Labat et al., 2005). Such increased rates of mutation and recombination events might also be the result of antibiotic treatment of the infant; leading to induction of the bacterial SOS response, which has been shown to increase mutation rates (Beaber et al., 2004; Michel, 2005).

Non-synonymous mutations were indeed detected in genes related to antibiotic tolerance, e.g., those involved in folate metabolism (*folA*—targeted by trimethoprim), fusaric acid resistance (*fusB*), ABC-transport and membrane permeability (porins; **Table S1**). Equivalently, the genomic deletion of the 54 kb region might have resulted from antibiotic mediated stress known to induce prophage excision and increase horizontal gene transfer in general (Beaber et al., 2004; Nanda et al., 2015).

Apart from small cryptic plasmids providing no obvious selective advantage to their bacterial host, gut-inhabiting *E. coli* isolates often carry plasmids that allow adaptation toward the human host by contributing virulence or antibiotic resistance factors (Johnson and Russo, 2005). In addition to the  $\beta$ -lactamase carrying pNK29 IncX plasmid, we identified a 114 kb plasmid (pNK29-2) in lineage A that was strikingly similar to other previously sequenced virulence plasmids from a diverse set of pathogenic *E. coli* strains (**Figure 3**). These plasmids have been shown to play a role in the initial stages of UTI infection in a mouse model in a different genetic host background, and could have provided lineage A with the necessary virulence factors leading to the successful UTI infection observed in the studied infant. Although it is generally believed that specific pathotypes of *E. coli* carry different virulence plasmids, plasmid backbones virtually identical to pNK29-2 have recently been found in *K. pneumoniae* as well as several divergent *E. coli* strains; with the earliest isolate dating back to 1974 (**Figure 3**). The high conservation of these plasmids suggests that they provide a

universal adaptive benefit to their ExPEC hosts regardless of infection site (Johnson and Nolan, 2009; Cusumano et al., 2010; Wijetunge et al., 2014).

When comparing the genetic composition of the currently sequenced virulence plasmids with high similarity to pNK29-2, it is clear that certain regions tend to preserve genetic features across host, geography and time (**Figures 3, 4**). These include mediators of iron acquisition, toxin production and putative copper resistance mediators (*scsC* and *scsD*; DebRoy et al., 2010). Carrying genes implicated in virulence, these plasmids could confer a survival advantage to their bacterial host during infection. Prior studies have examined the role of pUTI89 and pRS218 in urinary tract infections and neonatal meningitis, respectively (Cusumano et al., 2010; Wijetunge et al., 2014). These studies did not observe any phenotypic differences *in vitro* between the plasmid bearing and cured host in terms of growth rate, type 1 pilus expression or biofilm formation. However, they did observe a significant difference in infectivity using rodent infection models (Cusumano et al., 2010; Wijetunge et al., 2014).

A vital defense mechanism of the human body is to restrict iron from pathogens, thus acquisition and transport of iron is an important survival mechanism for ExPEC strains *in vivo* (Andrews et al., 2003). Therefore, iron acquisition could be beneficial for survival in many niches of the human body, including the densely populated gut microbiota, where access to iron is limited (Andrews et al., 2003; Nowrouzian et al., 2003). While Cusumano et al. speculated that the *cjr* operon of pUTI89 was beneficial in a UTI infection scenario due to its putative involvement in iron acquisition, we could not detect an advantage in neither absolute growth rate nor competitive fitness of lineage A carrying the pNK29-2 plasmid when grown in iron-limited conditions (Cusumano et al., 2010). Although the effect was small or non-existent in our experimental setups, pNK29-2 might provide an advantage by other means. For example, the pNK29-2 genes encoding siderophore receptors or transporters might provide an advantage, only if the available iron is on a certain form e.g., bound by its respective siderophore. As other iron acquisition systems are located on the chromosome of lineage A, the pNK29-2 encoded systems could be redundant in this strain, but might be selected in other hosts to encourage plasmid maintenance in a communal context.

Previous studies have shown that even minor differences in host genomes can be highly influential in determining plasmid establishment as well as subsequent adaptation and long term persistence (Humphrey et al., 2012; Porse et al., 2016). Thus, it is intriguing that virulence plasmids imposing a minor, but significant, fitness cost without providing any strongly selected phenotypes, can persist in a competitive environment such as the human gut. From our *in vivo* competition experiment, it is clear that the pNK29-2 carrying strain was less fit in the murine gut but does reach stable counts from day 7 to 12 (**Figure 5**). This stagnation in competition could encourage extended plasmid persistence and might be explained by changes in the growth rate of gut inhabiting strains over time; indicating a non-constant selection pattern

(Rang et al., 1999). Such selection patterns are known to occur in mixed bacterial populations encompassing social iron acquisition phenotypes and similar dynamics could take place in the competition between pNK29-2 carrying and plasmid-free strains in the gut (Stojiljkovic et al., 1993; Ross-gillespie et al., 2007). However, this does not seem to be the case from our growth rate measurements in iron-limiting conditions, where the plasmid carrying strain did not have an advantage on its own (Figure S3).

The measured fitness cost of the pNK29-2 in the lineage A strain was surprisingly similar between *in vitro* and *in vivo* setups, and is low compared to previous observations of plasmid costs (Vogwill and MacLean, 2015). While plasmids can in theory compensate their loss and fitness cost by re-infection of plasmid-free hosts (Slater et al., 2008), this is unlikely to be the case for pNK29-2 as we did not observe any transconjugants in our *in vitro* conjugation assays nor during the *in vivo* competition experiment. The inability of plasmids to conjugate might be explained by the fitness constraints a functional conjugation machinery can impose on certain hosts; supported by the existence of plasmid variants, such as pCE10A from Lu et al., lacking the conjugative transfer operon, that are otherwise identical to pNK29-2 (Lu et al., 2011; Porse et al., 2016).

Loss of pNK29-2 was not observed, neither *in vivo* nor *in vitro*, suggesting that primary stability mechanisms such as active segregation and toxin-antitoxin systems are the most important persistence parameters for these plasmids. This is consistent with the high degree of conservation of stability systems among all 14 plasmids examined here and further supported by the fact that we, as well as previous studies, have experienced considerable trouble curing strains from these plasmids; unless the *stbAB* operon is disrupted (Cusumano et al., 2010; Wijetunge et al., 2014).

By characterizing the genomes of persistent lineages of *E. coli* colonizing the gut of an infant, we observed substantial dynamics, highlighting that strains colonizing the human gut undergo continuous change. While genomic plasticity can lead to improved persistence, some elements are surprisingly stable. Our cost and stability characterizations suggest that a low cost and a high segregational stability, combined with plasmid-encoded universal virulence factors, presumed to increase fitness in a broad range of infection scenarios (Cusumano et al., 2010; Wijetunge et al., 2014), are likely the main parameters governing the success of endemic virulence plasmids. Further understanding of the factors contributing to genomic variation of gut colonizing pathogens will aid in rational interventions against the virulence and antibiotic resistance determinants widely disseminated among these isolates.

## DATA AVAILABILITY

All sequenced genomes can be accessed via the Bioproject PRJNA352659.

## AUTHOR CONTRIBUTIONS

NK, IA, and AW provided the *E. coli* isolates. AP did the *in vitro* work and strain tagging. JK performed the *in vivo* competition experiment. HG processed the sequencing data. AP, HG, and MS analyzed the sequencing data results and AP and HG wrote the manuscript with input from MS, JK, DA, NK, and IA.

## FUNDING

This research was funded by the EU H2020 ERC-20104-STG LimitMDR (638902) and the Danish Council for Independent Research Sapere Aude programme DFF -4004-00213, the Medical Faculty of the University of Göteborg (ALFGBG138401) and the Swedish Medical Research Council.

## ACKNOWLEDGMENTS

We thank Mari Cristina Rodriguez de Evgrafo and Marius Faza for aid in sequence library preparation. Christian Munck and Lejla Imamovic are thanked for helpful discussions. MS acknowledges support from the Novo Nordisk Foundation, and the Lundbeck Foundation.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fcimb.2017.00126/full#supplementary-material>

**Figure S1 | Plasmid map and BLAST comparison of pNK29.** pNK29 compared to IncX1 plasmids pOLA52 (outer ring, green) and pRPEC180\_47 (middle ring, blue). Open reading frames are drawn directionally (inner ring, black). Selected annotations are labeled outside the ring (see Table S2 for full annotation list).

**Figure S2 | Plasmid pNK29-3 compared to most similar plasmids in GeneBank.** The two ORFs encode putative replication and mobilization proteins and open reading frames are drawn directionally (inner ring, black).

**Figure S3 | Growth rate of lineage A with and without pNK29-2 under iron-limited conditions.**

**Table S1 | SNPs in different isolates of lineage A.** Table containing the SNPs from lineage A, including the annotation and whether the amino acid change was synonymous or non-synonymous. The isolates from days 2, 9, and 16 are left out as they did not have any SNP differences from the representative genome for the lineage (isolate taken at 2d).

**Table S2 | Annotations for pNK29.** Names in square brackets indicate homologs in pOLA52.

**Table S3 | SNPs in the pNK29-2 plasmid of lineage A compared to pUT189.** SNPs were identified by comparing the initial lineage A isolate (day 2) with pUT189 (CP000244.1).

**Table S4 | Plasmids similar to pNK29-2 and information on their origin.** The table presented in Figure 3, but more detailed information on size, location, and references. The GenBank ID refers to the plasmid sequence entry in GenBank and the Pubmed ID (PMID) to the study where the plasmid was initially described.

## REFERENCES

- Adlerberth, I., Svanborg, C., Carlsson, B., Mellander, L., Hanson, L. A., Jalil, F., et al. (1998). P fimbriae and other adhesins enhance intestinal persistence of *Escherichia coli* in early infancy. *Epidemiol. Infect.* 121, 599–608. doi: 10.1017/S0950268898001137
- Adlerberth, I., Lindberg, E., Aberg, N., Hesselmar, B., Saalman, R., Strannegård, I. L., et al. (2006). Reduced enterobacterial and increased staphylococcal colonization of the infantile bowel: an effect of hygienic lifestyle? *Pediatr. Res.* 59, 96–101. doi: 10.1203/01.pdr.0000191137.12774.b2
- Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., and Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. doi: 10.1186/1471-2164-12-402
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Andrews, S. C., Robinson, A. K., and Rodríguez-Quinones, F. (2003). Bacterial iron homeostasis. *FEMS Microbiol. Rev.* 27, 215–237. doi: 10.1016/S0168-6445(03)00055-X
- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75
- Beaber, J. W., Hochhut, B., and Waldor, M. K. (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* 427, 72–74. doi: 10.1038/nature02241
- Bidet, P., Burghoffer, B., Gautier, V., Brahimi, N., Mariani-Kurkdjian, P., El-Ghoneimi, A., et al. (2005). *In vivo* transfer of plasmid-encoded ACC-1 AmpC from *Klebsiella pneumoniae* to *Escherichia coli* in an infant and selection of impermeability to imipenem in *K. pneumoniae*. *Antimicrob. Agents Chemother.* 49, 3562–3565. doi: 10.1128/AAC.49.8.3562-3565.2005
- Brown, C. T., Sharon, I., Thomas, B. C., Castelle, C. J., Morowitz, M. J., and Banfield, J. F. (2013). Genome resolved analysis of a premature infant gut microbial community reveals a *Varibaculum cambriense* genome and a shift towards fermentation-based metabolism during the third week of life. *Microbiome* 1:30. doi: 10.1186/2049-2618-1-30
- Bruzskiewicz, E., Gottschalk, G., Ron, E., Hacker, J., and Dobrindt, U. (2009). Adaptation of pathogenic *E. coli* to various niches: genome flexibility is the key. *Microb. Pathog.* 6, 110–125. doi: 10.1159/000235766
- Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., et al. (2014). *In silico* detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* 58, 3895–3903. doi: 10.1128/AAC.02412-14
- Chen, S. L., Hung, C. S., Xu, J., Reigstad, C. S., Magrini, V., Sabo, A., et al. (2006). Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. *Proc. Natl. Acad. Sci. U.S.A.* 103, 5977–5982. doi: 10.1073/pnas.0600938103
- Chen, S. L., Wu, M., Henderson, J. P., Hooton, T. M., Hibbing, M. E., Hultgren, S. J., et al. (2013). Genomic diversity and fitness of *E. coli* strains recovered from the intestinal and urinary tracts of women with recurrent urinary tract infection. *Sci. Transl. Med.* 5, 184ra60. doi: 10.1126/scitranslmed.3005497
- Conlan, S., Thomas, P. J., Deming, C., Park, M., Lau, A. F., Dekker, J. P., et al. (2014). Single-molecule sequencing to track plasmid diversity of hospital-associated carbapenemase-producing Enterobacteriaceae. *Sci. Transl. Med.* 6, 254ra126. doi: 10.1126/scitranslmed.3009845
- Conlan, S., Park, M., Deming, C., Thomas, P. J., Young, A. C., Coleman, H., et al. (2016). Plasmid dynamics in KPC-positive *klebsiella pneumoniae* during long-term patient colonization. *MBio* 7, e00742-16. doi: 10.1128/mBio.00742-16
- Cusumano, C. K., Hung, C. S., Chen, S. L., and Hultgren, S. J. (2010). Virulence plasmid harbored by uropathogenic *Escherichia coli* functions in acute stages of pathogenesis. *Infect. Immun.* 78, 1457–1467. doi: 10.1128/IAI.01260-09
- DeRoy, C., Sidhu, M. S., Sarker, U., Jayarao, B. M., Stell, A. L., Bell, N. P., et al. (2010). Complete sequence of pEC14\_114, a highly conserved IncFIB/FIIA plasmid associated with uropathogenic *Escherichia coli* cystitis strains. *Plasmid* 63, 53–60. doi: 10.1016/j.plasmid.2009.10.003
- Diard, M., Garry, L., Selva, M., Mosser, T., Denamur, E., and Matic, I. (2010). Pathogenicity-associated islands in extraintestinal pathogenic *Escherichia coli* are fitness elements involved in intestinal colonization. *J. Bacteriol.* 192, 4885–4893. doi: 10.1128/JB.00804-10
- Dias, R. C., Moreira, B. M., and Riley, L. W. (2010). Use of fimH single-nucleotide polymorphisms for strain typing of clinical isolates of *Escherichia coli* for epidemiologic investigation. *J. Clin. Microbiol.* 48, 483–488. doi: 10.1128/JCM.01858-09
- Drasar, B. S., and Hill M. J. (1974). *Human Intestinal Flora*. London: Academic Press Inc.
- Foxman, B. (2010). The epidemiology of urinary tract infection. *Nat. Rev. Urol.* 7, 653–660. doi: 10.1038/nrurol.2010.190
- Goren, M. G., Carmeli, Y., Schwaber, M. J., Chmelnitsky, I., Schechner, V., and Navon-Venezia, S. (2010). Transfer of carbapenem-resistant plasmid from *Klebsiella pneumoniae* ST258 to *Escherichia coli* in patient. *Emerg. Infect. Dis.* 16, 1014–1017. doi: 10.3201/eid1606.091671
- Guynet, C., Cuevas, A., Moncalián, G., and de la Cruz, F. (2011). The stb operon balances the requirements for vegetative stability and conjugative transfer of plasmid R388. *PLoS Genet.* 7:e1002073. doi: 10.1371/journal.pgen.1002073
- Humphrey, B., Thomson, N. R., Thomas, C. M., Brooks, K., Sanders, M., Delsol, A. A., et al. (2012). Fitness of *Escherichia coli* strains carrying expressed and partially silent IncN and IncP1 plasmids. *BMC Microbiol.* 12:53. doi: 10.1186/1471-2180-12-53
- Johnson, T. J., and Nolan, L. K. (2009). Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiol. Mol. Biol. Rev.* 73, 750–774. doi: 10.1128/MMBR.00015-09
- Johnson, J. R., and Russo, T. A. (2005). Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. *Int. J. Med. Microbiol.* 295, 383–404. doi: 10.1016/j.ijmm.2005.07.005
- Johnson, T. J., Danzeisen, J. L., Youmans, B., Case, K., Llop, K., Munoz-Aguayo, J., et al. (2016). Separate F-Type plasmids have shaped the evolution of the H 30 subclone of *Escherichia coli* sequence type 131. *mSphere* 1, e00121–e00116. doi: 10.1128/mSphere.00121-16
- Karami, N., Martner, A., Enne, V. I., Swerkersson, S., Adlerberth, I., and Wold, A. E. (2007). Transfer of an ampicillin resistance gene between two *Escherichia coli* strains in the bowel microbiota of an infant treated with antibiotics. *J. Antimicrob. Chemother.* 60, 1142–1145. doi: 10.1093/jac/dkm327
- Köhler, C. D., and Dobrindt, U. (2011). What defines extraintestinal pathogenic *Escherichia coli*? *Int. J. Med. Microbiol.* 301, 642–647. doi: 10.1016/j.ijmm.2011.09.006
- Kuhlman, T. E., and Cox, E. C. (2010). Site-specific chromosomal integration of large synthetic constructs. *Nucleic Acids Res.* 38:e92. doi: 10.1093/nar/gkp1193
- Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12
- Labat, F., Pradillon, O., Garry, L., Peuchmaur, M., Fantin, B., and Denamur, E. (2005). Mutator phenotype confers advantage in *Escherichia coli* chronic urinary tract infection pathogenesis. *FEMS Immunol. Med. Microbiol.* 44, 317–321. doi: 10.1016/j.femsim.2005.01.003
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi: 10.1186/gb-2009-10-3-r25
- Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., et al. (2012). Multilocus sequence typing of total-genome-sequenced bacteria. *J. Clin. Microbiol.* 50, 1355–1361. doi: 10.1128/JCM.06094-11
- Lasaro, M., Liu, Z., Bishar, R., Kelly, K., Chattopadhyay, S., Paul, S., et al. (2014). *Escherichia coli* isolate for studying colonization of the mouse intestine and its application to two-component signaling knockouts. *J. Bacteriol.* 196, 1723–1732. doi: 10.1128/JB.01296-13
- Lee, S. M., Wyse, A., Leshar, A., Everett, M. L., Lou, L., Holzkecht, Z. E., et al. (2010). Adaptation in a mouse colony monoassociated with *Escherichia coli* K-12 for more than 1,000 Days. *Appl. Environ. Microbiol.* 76, 4655–4663. doi: 10.1128/AEM.00358-10
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Lu, S., Zhang, X., Zhu, Y., Kim, K. S., Yang, J., and Jin, Q. (2011). Complete genome sequence of the neonatal-meningitis-associated *Escherichia coli* strain CE10. *J. Bacteriol.* 193, 7005. doi: 10.1128/JB.06284-11



- Lutz, R., and Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/11-12 regulatory elements. *Nucleic Acids Res.* 25, 1203–1210.
- M9 minimal medium (standard) (2010). M9 minimal medium (standard). *Cold Spring Harb. Protoc.* 2010:pdh.rec12295. doi: 10.1101/pdb.rec12295
- Marchesi, J. R., Adams, D. H., Fava, F., Hermes, G. D., Hirschfield, G. M., Hold, G., et al. (2015). The gut microbiota and host health: a new clinical frontier. *Gut* 65, 330–339. doi: 10.1136/gutjnl-2015-309990
- Michel, B. (2005). After 30 years of study, the bacterial SOS response still surprises us. *PLoS Biol.* 3, 1174–1176. doi: 10.1371/journal.pbio.0030255
- Morowitz, M. J., Denef, V. J., Costello, E. K., Thomas, B. C., Poroyko, V., Relman, D. A., et al. (2011). Strain-resolved community genomic analysis of gut microbial colonization in a premature infant. *Proc. Natl. Acad. Sci. U.S.A.* 108, 1128–1133. doi: 10.1073/pnas.1010992108
- Mulvey, M. A., Schilling, J. D., and Hultgren, S. J. (2001). Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infect. Immun.* 69, 4572–4579. doi: 10.1128/IAI.69.7.4572-4579.2001
- Myhrvold, C., Kotula, J. W., Hicks, W. M., Conway, N. J., and Silver, P. A. (2015). A distributed cell division counter reveals growth dynamics in the gut microbiota. *Nat. Commun.* 6:10039. doi: 10.1038/ncomms10039
- Nanda, A. M., Thormann, K., and Frunzke, J. (2015). Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. *J. Bacteriol.* 197, 410–419. doi: 10.1128/JB.02230-14
- Norman, A., Hansen, L. H., She, Q., and Sørensen, S. J. (2008). Nucleotide sequence of pOLA52: a conjugative IncX1 plasmid from *Escherichia coli* which enables biofilm formation and multidrug efflux. *Plasmid* 60, 59–74. doi: 10.1016/j.plasmid.2008.03.003
- Norman, A., Hansen, L. H., and Sørensen, S. J. (2009). Conjugative plasmids: vessels of the communal gene pool. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 2275–2289. doi: 10.1098/rstb.2009.0037
- Nowrouzian, F., Hesselmar, B., Saalman, R., Strannegård, I. L., Aberg, N., Wold, A. E., et al. (2003). *Escherichia coli* in infants' intestinal microflora: colonization rate, strain turnover, and virulence gene carriage. *Pediatr. Res.* 54, 8–14. doi: 10.1203/01.PDR.0000069843.20655.EE
- Nowrouzian, F. L., Wold, A. E., and Adlerberth, I. (2005). *Escherichia coli* strains belonging to phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants. *J. Infect. Dis.* 191, 1078–1083. doi: 10.1086/427996
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T., et al. (2015). Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31, 3691–3693. doi: 10.1093/bioinformatics/btv421
- Porse, A., Schoønning, K., Munck, C., and Sommer, M. O. (2016). Survival and evolution of a large multidrug resistance plasmid in new clinical bacterial hosts. *Mol. Biol. Evol.* 33, 2860–2873. doi: 10.1093/molbev/msw163
- Quinlan, A. R., and Hall, I. M. (2010). Bed tools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Rang, C. U., Licht, T. R., Midtvedt, T., Conway, P. L., Chao, L., Krogfelt, K. A., et al. (1999). Estimation of growth rates of *Escherichia coli* BJ4 in streptomycin-treated and previously germfree mice by in situ rRNA hybridization. *Clin. Diagn. Lab. Immunol.* 6, 434–436. doi: 10.1128/JB.01581-07
- Ross-gillespie, A., Gardner, A., West, S. A., and Griffin, A. S. (2007). Frequency dependence and cooperation: theory and a test with bacteria. *Am. Nat.* 170, 331–342. doi: 10.1086/519860
- Salyers, A. A., Gupta, A., and Wang, Y. (2004). Human intestinal bacteria as reservoirs for antibiotic resistance genes. *Trends Microbiol.* 12, 412–416. doi: 10.1016/j.tim.2004.07.004
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Sharon, I., Morowitz, M. J., Thomas, B. C., Costello, E. K., Relman, D. A., and Banfield, J. F. (2013). Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res.* 23, 111–120. doi: 10.1101/gr.142315.112
- Simonsen, L. (1991). The existence conditions for bacterial plasmids: theory and reality. *Microb. Ecol.* 22, 187–205.
- Slater, F. R., Bailey, M. J., Tett, A. J., and Turner, S. L. (2008). Progress towards understanding the fate of plasmids in bacterial communities. *FEMS Microbiol. Ecol.* 66, 3–13. doi: 10.1111/j.1574-6941.2008.00505.x
- Smajs, D., and Weinstock, G. M. (2001). The iron- and temperature-regulated cjrBC genes of *Shigella* and enteroinvasive *Escherichia coli* strains code for colicin J5 uptake. *J. Bacteriol.* 183, 3958–3966. doi: 10.1128/JB.183.13.3958-3966.2001
- Smith, J. L., Fratamico, P. M., and Gunther, N. W. (2007). Extraintestinal pathogenic *Escherichia coli*. *Foodborne Pathog. Dis.* 4, 134–163. doi: 10.1089/fpd.2007.0087
- Smorawska, M., Szuplewska, M., Zaleski, P., Wawrzyniak, P., Maj, A., Plucienniczak, A., et al. (2012). Mobilizable narrow host range plasmids as natural suicide vectors enabling horizontal gene transfer among distantly related bacterial species. *FEMS Microbiol. Lett.* 326, 76–82. doi: 10.1111/j.1574-6968.2011.02432.x
- Sommer, M. O., Church, G. M., and Dantas, G. (2010). The human microbiome harbors a diverse reservoir of antibiotic resistance genes. *Virulence* 1, 299–303. doi: 10.4161/viru.1.4.12010
- Stamatakis, A. (2014). RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Stoesser, N., Sheppard, A. E., Pankhurst, L., De Maio, N., Moore, C. E., Sebra, R., et al. (2016). Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *MBio* 7, e02162–e02115. doi: 10.1128/mBio.02162-15
- Stojiljkovic, I., Cobeljic, M., and Hantke, K. (1993). *Escherichia coli* K-12 ferrous iron-uptake mutants are impaired in their ability to colonize the mouse intestine. *FEMS Microbiol. Lett.* 108, 111–115.
- Sullivan, M. J., Petty, N. K., and Beatson, S. A. (2011). Easyfig: a genome comparison visualizer. - PubMed - NCBI. *Bioinformatics* 27, 1009–1010. doi: 10.1093/bioinformatics/btr039
- de Toro, M., Garcillán-barcia, M. P., and De La Cruz, F. (2014). Plasmid diversity and adaptation analyzed by massive sequencing of *Escherichia coli* plasmids. *Microbiol. Spectr.* 2, 1–16. doi: 10.1128/microbiolspec.PLAS-0031-2014
- Trobos, M., Lester, C. H., Olsen, J. E., Frimodt-Møller, N., and Hammerum, A. M. (2009). Natural transfer of sulphonamide and ampicillin resistance between *Escherichia coli* residing in the human intestine. *J. Antimicrob. Chemother.* 63, 80–86. doi: 10.1093/jac/dkn437
- Vogwill, T., and MacLean, R. C. (2015). The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. *Evol. Appl.* 8, 284–295. doi: 10.1111/eva.12202
- Wang, C. T. (1996). Concentration of arsenic, selenium, zinc, iron and copper in the urine of blackfoot disease patients at different clinical stages. *Eur. J. Clin. Chem. Clin. Biochem.* 34, 493–497.
- Wijetunge, D. S., Karunathilake, K. H., Chaudhari, A., Katani, R., Dudley, E. G., Kapur, V., et al. (2014). Complete nucleotide sequence of pRS218, a large virulence plasmid, that augments pathogenic potential of meningitis-associated *Escherichia coli* strain RS218. *BMC Microbiol.* 14:203. doi: 10.1186/s12866-014-0203-9
- Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107
- Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., and Wishart, D. S. (2011). PHAST: a fast phage search tool. *Nucleic Acids Res.* 39, 347–352. doi: 10.1093/nar/gkr485

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Porse, Gumpert, Kubicek-Sutherland, Karami, Adlerberth, Wold, Andersson and Sommer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.